

미세먼지의 빅데이터/AI 분석 및 예측을 위한 IoT 측정 단말기 개발

우동식, 백봉현*

대구가톨릭대학교, *(주)아르고스

dswoo@cu.ac.kr, *wefbbh@argosinc.co.kr

Development of a IoT measurement terminal for big data/AI analysis and prediction of fine dust

Woo Dong Sik, Back Bong-Hyun*

Daegu Catholic Univ., *Argos, Inc.

요약

본 논문에서는 지역별로 상이한 환경조건(산업, 산림, 인구밀도, 자동차 등)과 다양한 산업 분야에서 발생하는 환경변수 데이터와 지역(구군별)별 발생하는 실제 대기 정보(미세먼지)를 빅데이터 및 인공지능 기법을 이용하여 미세먼지의 원인 분석과 예측이 가능한 빅데이터 플랫폼에 사용가능한 태양광을 이용한 IoT 측정 단말기를 개발하였다. 이를 통하여 최근 이슈화 되고 있는 미세먼지의 보다 정확한 세부 지역별 예측에 있어 빅데이터와 인공지능 기술을 활용한 플랫폼 서비스에 활용 가능성을 보였다.

I. 서론

최근 대기환경오염의 위험에 대한 환경 문제가 이슈화되고 있으며, 특히 잦은 초미세 먼지의 발생으로 인해 유해성 문제가 점차 대두되고 있다. 2016년부터 정부에서는 경유차규제와 석탄화력 발전소 규제를 통한 저감 정책을 발표하였으나, 정확한 미세먼지의 발생원인과 해결책을 제시하지 못하고 있다. 대기환경에 영향을 미치는 지역별 환경변수의 차이로 정확한 미세먼지 등의 예측정보 제공의 어려움이 있다[1-2]. 따라서 지역적 환경조건에 기반하여 공공데이터기반 환경변수를 위한 빅데이터 처리 플랫폼과 상세 지역별 대기질 정보 및 데이터를 제공하기 위한 통계분석 및 인공지능 기술기반 예측시스템을 제안하고 이러한 환경데이터 수집을 위한 IoT 측정 단말기를 제안하였다.

II. 본론

기존의 기상청 데이터를 기반으로 하는 전국단위의 미세먼지 측정예보 방식의 신뢰성 부족을 감안하여 친환경 대기질 측정 장비로부터의 상세 지역별 대기질 데이터 및 공공데이터 측정하도록 하였다. 지역별 환경변수 특성과 기상청 데이터를 결합한 빅데이터 수집·저장·분석이 가능한 플랫폼을 제안하였다. 지역 또는 소규모 단위의 보다 정확한 미세먼지의 예측을 위해 그라디언트 부스팅기법의 인공지능기법을 적용하여 상세 지역별 대기질정보 및 각 지역별 정확한 미세먼지 예측을 통한 대국민 생활안전 서비스를 제공할 수 있다. 이를 위해 로라(LoRa)망을 통한 태양광발전 소규모 미세먼지 측정기로부터 수집된 기상데이터(미세먼지, 초미세먼지, 풍향, 이산화탄소 등)를 구간별/지역별로 수집하여 빅데이터분석 처리를 통해 구간별/지역별 상세한 대기질맵(지도)을 구축하도록 하고, 외부데이터(기상청, 국토지리정보, 시도별 자동차 보유량 및 지역별 공장, 농업 등의 생산환경)등의 각 기관별로 제공되는 다양한 데이터포맷의 지역 환경변수를 연계하여 지역의 미세먼지 발생원인과 예측할 수 있는 빅데이터

수집·저장·분석시스템과 시각화 서비스를 개발하였다. 제안된 플랫폼의 구성도는 그림 1과 같다. 웹 UI를 통한 데이터의 수집·관리, DataBase와 통계분석도구 R을 활용하여 빅데이터 통계분석 및 시각화 및 자동 레포팅을 위한 자동화 통계가 가능하도록 하였다.

1) 데이터 수집

미세먼지 예측을 위해 대기질 측정 Device로부터 대기환경을 측정하여 대기 및 고도·위치 등의 정보를 소프트웨어로 로딩을 통해 통합 수집모듈로 송신하는 네트워크 모듈을 개발하였다. 공공데이터 수집 모듈은 지역의 정확한 대기질정보와 대기질 예측 및 발생원인을 분석하기 위해 기상청, 환경관리공단 및 통계청 등에서 측정된 데이터를 수집할 수 있도록 하였다. 통합데이터 수집 모듈은 Device 및 공공데이터에서 수집된 데이터의 포맷이 상이함에 따라 분석에 필요한 데이터 포맷 및 수집경로를 관리할 수 있도록 한다. 그림2는 제안된 태양광을 이용한 LTE IoT 측정 단말기의 사진이다. LTE Cat. M1 통신 모듈을 적재하고, 아두이노보드와 미세먼지 측정 모듈 연동을 하였다. 태양광발전패널과 내부배터리를 연동하여 사용하고 TCP/IP 네트워크를 통한 데이터 수신이 가능하도록 하였다.

2) 데이터 저장

수집된 대량의 데이터의 안정적, 신뢰적 저장을 위해 HDFS(Hadoop Distributed File System)시스템을 구축하고, 분석된 결과에 대한 빠른 저장과 조회를 위해 NoSQL(MongoDB)를 활용하도록 한다.

3) 데이터 분석

수집된 데이터로부터 대기오염에 대한 고차원 통계분석을 위해 오픈소스 기반의 통계R을 웹버전으로 개발 한다. 또한 스케줄링에 의해 지속적으로 수집된 데이터의 통계분석 결과를 레포팅할 수 있도록 하였다. 대기오염 예측분석은 각 지역별 Device로부터 수집된 대기정보와 공공데이터로부터 수집된 공장 등의 지역현황 정보 및 기상청으로 부터의 전국 기상데이터 등을 기반으로 각 지역별(시군구동)의 상세한 미세먼지 현황과 예

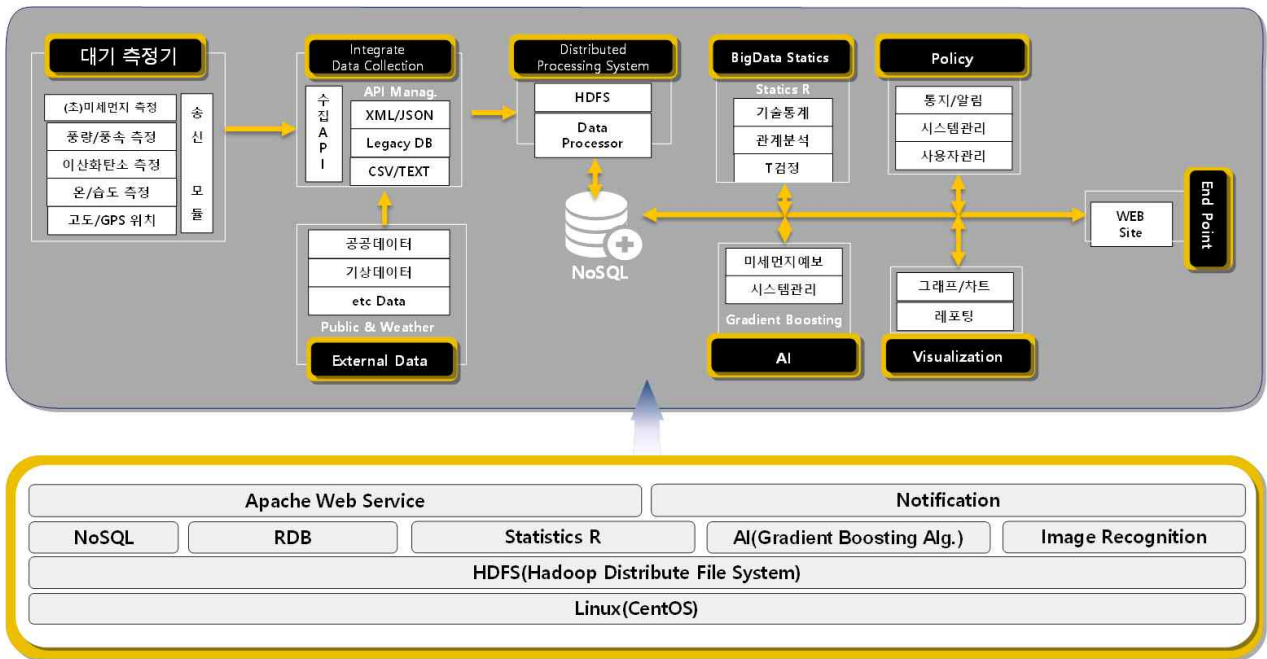


그림1. 제안된 플랫폼 구성도

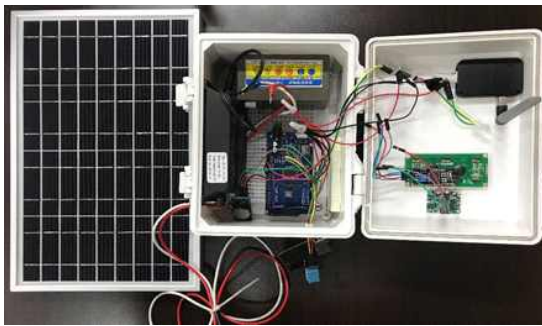


그림2. 제작된 LTE기반 IoT 측정 단말기

측을 통한 우리동네 미세먼지 정보를 산출하며, 수집된 데이터를 AI기법 중 그래디언트 부스팅 알고리즘(Gradient Boosting Algorithm)을 적용하여 대기질 분석 및 예측을 하였다.

4) 시각화 및 데이터제공

분석된 결과물의 시각화 전달을 위해 웹에 미세먼지의 현황을 실시간으로 조회할 수 있는 웹사이트를 활용하며 회원관리, 외부데이터 연계 및 정책 등을 설정할 수 있도록 백오피스를 활용할 수 있다. 원본데이터 및 빅데이터 분석정보를 CSV, XML, JSON 형식으로 데이터를 제공할 수 있도록 API를 제공할 수 있다. 그림 3은 에어코리아 API와 연동하여 홈/미세먼지 농도를 PM10, PM2.5로 구분하여 시간대별로 노출하여 최종 수집 시간의 선택 지역의 단일 미세먼지 노출 사례를 보여주고 있다.

III. 결론

본 논문에서는 기존의 기상청 데이터를 기반으로 하는 전국단위의 미세먼지 측정예보 방식에서의 고려되지 못한 지역별환경변수 특성과 기상청 데이터를 결합한 빅데이터 수집·저장·분석 플랫폼을 위한 IoT 측정 단말기를 개발하였다. 개발된 측정 단말기로 수집된 데이터를 활용하여 지역단위의 보다 정확한 미세먼지의 예측을 위해 그래디언트 부스팅 기법의 인공지능기법을 적용하여 정확한 미세먼지 예측이 가능하도록 하였다.



그림3. 홈/미세먼지 AI Air API 사례

제안된 시스템은 세분화된 대기질 정보분석과 예측 시스템 등 다양하게 응용될 수 있을 것이다.

ACKNOWLEDGMENT

본 논문은 (주)아르고스와 산학공동연구에 의하여 수행되었음.

참 고 문 헌

- [1] 강성원의, “환경 빅데이터 분석 및 서비스 개발,” 한국환경정책·평가연구원, 2017.
- [2] 임준묵, “기상환경데이터와 머신러닝을 활용한 미세먼지농도 예측 모델,” *한국IT서비스학회지* 제18권, 제1호, pp. 173-186, 2019.